

INTERACTIVE ARTICLE COVER

The Role of Congruency in Collocation Acquisition: A Case Study of Vietnamese Students Learning English Adjective + Noun Collocations

RUPKATHA JOURNAL
ON INTERDISCIPLINARY STUDIES IN HUMANITIES



About the Journal

Journal DOI	https://dx.doi.org/10.21659/rupkatha
Journal Home	www.rupkatha.com
Indexed by	Scopus Web of Science: ESCI DOAJ

About the Issue

Themed issue	Volume 15, number 1, 2023 Current and Future Directions in TESOL Studies
Edited by	Dr John R. Baker
Issue DOI	https://doi.org/10.21659/rupkatha.v15n1
TOC	https://rupkatha.com/v15n1.php

About the Article

Title	The Role of Congruency in Collocation Acquisition: A Case Study of Vietnamese Students Learning English Adjective + Noun Collocations	
Author/s	Ngoc Thai Bao Pham	
Affiliation	University of Social Sciences and Humanities, Vietnam National University, Ho Chi Minh City, Vietnam	
Author ID	0000-0001-5070-678X	
DOI	https://doi.org/10.21659/rupkatha.v15n1.05	Pages: 1-15
Abstract	https://rupkatha.com/v15n105	
Full-text PDF	https://rupkatha.com/V15/n1/v15n105.pdf	
Article History	First Published: 06 May 2023	
Article Impact	Check Dynamic Impact	
Copyright	Aesthetics Media Services	

The Role of Congruency in Collocation Acquisition: A Case Study of Vietnamese Students Learning English Adjective + Noun Collocations

Ngoc Thai Bao Pham

University of Social Sciences and Humanities, Vietnam National University, Ho Chi Minh City, Vietnam. ORCID: 0000-0001-5070-678X.

Email: phamthaibaongoc@hcmussh.edu.vn

Abstract

Collocational competence is of critical importance to EFL learners' language proficiency. Due to limited research on L1 influence on L2 learners' development of receptive and productive knowledge of adjective + noun collocations, especially in the context of Vietnam, this study is an attempt to address these gaps by investigating whether Vietnamese learners could recognize and use congruent English adjective + noun collocations more accurately than incongruent ones. Data were collected from 72 English-major undergraduates in a Vietnamese university who completed three types of collocational tests: the Receptive Knowledge Test, the Productive Knowledge Test, and the paragraph-writing test. Results showed that, in contrast to language transfer theory, congruent collocations posed considerably more problems for the students in identifying the well-formed lexical combinations and using them accurately in writing. These findings highlight the importance of explicitly teaching congruent and incongruent collocations in EFL classrooms.

Keywords: congruency, language transfer, congruent collocation, incongruent collocation, adjective + noun collocation

1. Introduction

Vocabulary typically functions not as single words linked with each other by grammar but rather as multi-word phraseological units (Foster, 2001; Schmitt, 2012). According to Erman and Warren (2000), approximately fifty percent of spoken and written English discourse is formulaic sequences, i.e., prefabricated chunks or conventionalized word combinations. Among various types of formulaic language, collocations are regarded as "the most powerful force in the creation and comprehension of all naturally-occurring text" (Lewis, 2000, p. 53). Due to the importance of formulaic language in general and collocations in particular in learning English, the lack of collocational knowledge can create barriers to EFL learners' effective communication and language proficiency (Celce-Murcia, 2008; Wray, 2002).

Despite its importance, collocations cause considerable difficulty for EFL learners, even those at advanced levels (Laufer & Waldman, 2011; Nesselhauf, 2003). Recent studies have shown that L1 influence is one of the major causes of collocation errors (Chorbwhan & McLellan, 2016; Laufer & Waldman, 2011; Wolter, 2006; Yamashita & Jiang, 2010). In terms of L1-L2 collocation congruency, a great deal of research has found that L2 learners tend to perform better on congruent

collocations than incongruent ones, especially when making lexical choices (Nesselhauf, 2003; Patekar & Košuta, 2022; Wolter & Gyllstad, 2011; Yamashita & Jiang, 2010), whereas other studies have shown that it is congruent collocations that are more likely to be troublesome to L2 learners in their writing (Cao & Badger, 2021; Men, 2018). The contradictory results from earlier research appear to increase the complexity and uncertainty of this issue rather than effectively resolve it. Besides, most studies mentioned earlier, such as Nesselhauf (2003), Wolter and Gyllstad (2011), Cao and Badger (2021), and Patekar and Košuta (2022), focused on verb + noun collocations rather than other types, for example, adjective + noun collocations, which are also one of the most commonly used by native speakers (Benson et al., 2010) and one of the most challenging to L2 students (Gui & Yang, 2003). In addition, there has been little research on L2 learners' receptive and productive collocational knowledge concerning L1-L2 congruency, especially in Vietnam. These are the reasons why this study is aimed at investigating whether Vietnamese learners perform better with congruent collocations than incongruent ones in terms of recognizing and producing conventional adjective + noun collocations. This aim can be addressed through the following research questions:

[1] Do Vietnamese learners recognize congruent conventional adjective + noun collocations significantly more accurately than incongruent ones?

[2] Do Vietnamese learners produce congruent conventional adjective + noun collocations significantly more accurately than incongruent ones?

Within the scope of this study, *receptive knowledge* refers to the ability to recognize or identify appropriate adjective + noun combinations in writing, not in listening, and *productive knowledge* refers to the ability to recall and use adjective + noun collocations correctly in writing, not in speaking. The term *conventional collocations* is used to imply that these combinations are frequently used by native speakers and therefore considered correct or well-formed collocations (in comparison with *unconventional collocations*, i.e., incorrect or deviant collocations produced by L2 learners).

2. Literature Review

2.1 Definitions of collocation

An adequate definition of collocation with clear criteria is of great importance in order to distinguish collocations from other linguistic expressions. Three major approaches frequently adopted by researchers when defining collocation are the psychological, frequency-based, and phraseological approaches. While the psychological approach focuses on the associative bonds between words stored in the mental lexicon of native speakers (Hoey, 2005; Leech, 1974), the frequency-based approach, also known as the statistically-based or the Firthian approach, is concerned with the syntagmatic relations between lexical items within a set span in texts and their frequently repeated co-occurrences to the extent of statistical significance (Clear, 1993; Firth, 1957; Sinclair, 1991). Unlike the first two approaches, the phraseological approach concentrates primarily on the criteria for identifying collocations, including the commutability and the specialised senses of one element of a collocation as well as its semantic transparency in order to differentiate a collocation with a free combination and an idiom (Cowie, 1981; Howarth, 1998;

Laufer, 2010). The frequency-based and phraseological approaches are highly useful for this research, which involves a set of criteria for identifying collocations in EFL students' writings before analyzing their L1-induced errors. Because each approach has its advantages and disadvantages, it is recommended that a definition of collocation should incorporate elements from both the frequency-based approach and the phraseological approach (Cao & Deignan, 2019; Granger, 2018; Kurosaki, 2012).

However, without clear lines of demarcation between collocation and other linguistic expressions, even when the two approaches are employed to define collocation, it is still very difficult to decide whether a word combination is a restricted collocation or, for instance, a free combination. This is why two specific criteria, frequency (FREQ) and mutual information (MI), are recommended for corpus-based research on collocation. According to Wolter and Yamashita (2015), a minimum MI-score threshold of 3 is frequently used to indicate significant co-occurrence between two lexical items, showing how closely related these words are (p. 1201). Regarding frequency, Hong, Rahim, Hua, and Salehuddin (2011) specify that the minimum FREQ of 5 tokens in the British National Corpus (BNC) proves sufficient to be the standard threshold for a word combination to be considered a collocation. Meanwhile, Shin and Chon (2019) point out that 20 repetitions or above should be the minimum cut-off point for frequency in the Corpus of Contemporary American English (COCA) due to its relatively large size of more than one billion words, which is ten times as many words as in BNC (p. 611). Since this research used COCA as a reference source to identify collocations, the operational definition of collocation is constructed as shown below:

A collocation is a combination of two or more words which (1) is relatively transparent in meaning, (2) has a restricted range of co-occurrence in at least one of their elements, (3) presents a minimum MI score of 3.0, and (4) has a minimum FREQ of 20 tokens in COCA.

2.2 Collocation congruency

Collocation congruency between L1 and L2 is typically discussed in terms of translation equivalence (Men, 2018; Nesselhauf, 2005; Wolter & Gyllstad, 2011). Nesselhauf (2005) defines *congruency* as the presence or absence of a literal translation equivalent. A collocation is considered congruent if its constituent words in one language have direct translation equivalence, i.e., the kind of equivalence attained by literal word-for-word translation in another language. For instance, *strong wind* in English can be translated word-for-word as *gió mạnh* in Vietnamese. In contrast, an incongruent collocation refers to an equivalent combination that does not have equivalent components in the two languages when these components are rendered separately from one language to another. An example of incongruent collocations is *soft drink* in English with its literal Vietnamese translation as **nước mềm*, which is infelicitous.

According to Nesselhauf (2005), the concept of congruency is difficult to understand adequately. The problem is that it is not always clear whether or not two words are direct translation equivalents of each other in two languages. For instance, *strong tea* in English is *trà đặc* in Vietnamese. However, *trà đặc* is literally **thick tea* when translated word-for-word from Vietnamese into English, but it is still uncertain whether *strong tea* has direct translation equivalence with *trà đặc*. To address this problem, previous studies adopted the strategy of back translation (Men, 2018; Nesselhauf, 2005), where the translated word in question is rendered back word-for-word from the target language (e.g., *trà đặc* in Vietnamese) into the source language

(e.g., **thick tea* in English) out of context. If this word combination (e.g., **thick tea*) is infelicitous, the collocation in question (e.g., *strong tea*) is considered incongruent; otherwise, it is a congruent collocation. To minimize the researcher's subjective judgment on collocation congruency, Men (2018) suggested using reliable bilingual dictionaries during the back-translation process.

2.3 Recent studies related to congruency in collocation learning

Generally, most researchers have agreed that EFL learners tend to encounter more difficulties with incongruent collocations than congruent ones, so it is incongruent collocations that need relatively special attention in the collocation teaching and learning processes (Jiang, 2022; Nesselhauf, 2003; Peters, 2015; Wolter & Gyllstad, 2011, 2013; Yamashita & Jiang, 2010; Yigit, 2021). For example, employing a primed lexical decision task and a test of receptive collocational knowledge (2011), and an acceptability judgment task (2013), Wolter and Gyllstad pointed out that incongruent collocations caused considerably more errors for non-native English speakers (L1 Swedish), even at the advanced level of English, than congruent collocations. This view is supported by Yamashita and Jiang (2010), who compared the performance of Japanese participants at different levels of English proficiency on a phrase-acceptability judgment task. The findings revealed that both EFL and ESL learners dealt with congruent collocations far better than incongruent ones, though ESL learners outperformed their EFL counterparts in terms of accuracy. Along the same lines, Yigit (2021) analyzed the results from a phrase-judgment test and concluded that Turkish students, irrespective of their level of English vocabulary, tend to have lower error rates and shorter reaction times with congruent collocations in comparison with incongruent ones. Similarly, Peters (2015) and Jiang (2022) concluded that incongruent collocations appear to be more difficult for EFL learners (L1 Dutch and L1 Chinese irrespectively) to recollect compared with congruent ones, though different tests were used in their studies: Peters (2015) used the two form recall tests, whereas Jiang (2022) used an acceptability judgment task.

In contrast to the studies mentioned above, Men (2018), El-Dakhs, Salem, and Al-Haqbani (2020), and Cao and Badger (2021) reported that the participants in their studies made more mistakes when using congruent collocations, which are commonly believed to be less troublesome to L2 learners compared with incongruent ones. Rather than measuring the learner's receptive knowledge or their form retention of collocation via different tests, Men (2018) made use of the Chinese Learner English Corpus to examine how Chinese learners of English at different proficiency levels of English used congruent and incongruent collocations in their writings. In the same vein, Cao and Badger (2021) identified and analyzed collocation errors in a Vietnamese-speaking learners' database of argumentative essays, concluding that unconventional collocations result more from congruent collocations. However, as several students tend to avoid using collocations that they are uncertain about (perhaps to minimize possible mistakes and therefore gain higher scores), or they simply do not need to use certain collocations which are not suitable to the writing topics, it appears to be inadequate if researchers use students' writings as the only resource to evaluate their productive knowledge of collocations.

In contrast to Men (2018) and Cao and Badger (2021), El-Dakhs, Salem, and Al-Haqbani (2020) did not measure learners' collocational knowledge through their writings but through their answers to objective test items, including a multiple-choice task to evaluate their form recognition and a gap-fill task to test their form recall. The participants were Arab freshmen and seniors. The findings

revealed no statistically significant difference in students' form recognition of congruent and incongruent collocations, except in the case of last-year students, in contrast to a significant difference in terms of form recall irrespective of collocational congruency and years of study. Although using objective test types enables researchers to bring to the tests challenging collocations that students tend to avoid in speaking or writing, it does not provide information about learners' tendency to choose a specific type of collocation or their actual use of collocations to express their ideas in longer discourse. Therefore, it is necessary for a researcher to combine both objective (e.g., multiple choice questions) and subjective (e.g., essay writing) test items when designing research instruments to gain a more comprehensive understanding of L2 learners' recognition and production of collocations.

In brief, the review of the literature reveals that there has been an insufficient number of studies whose research instruments include both objective and subjective test items for a deeper insight into L2 students' receptive and productive collocational knowledge, not to mention that L2 learners' productive knowledge of collocations appears to be inadequately assessed. In addition, the contradictory findings of previous studies concerning the role of congruency in collocation acquisition appear to exacerbate the issue rather than resolve it. Therefore, given the inconsistent findings about collocation congruency and the lack of research on this issue, especially in the context of Vietnam, this study is an attempt to address these gaps in the literature by examining whether Vietnamese learners perform better with congruent collocations than incongruent ones in terms of recognizing and producing conventionalized collocations, specifically adjective + noun collocations.

3. Methodology

3.1 Participants

Seventy-two English-major sophomores, aged 20-21, at a university in Vietnam, were recruited from three intact classes using convenience sampling. At the time of this research, the students were at the intermediate or upper-intermediate level (equivalent to B1 and B2 CEFR) and were taking a Grammar course in which word combinations, including collocations, were part of the syllabus.

3.2 Research instruments

This study used two main research instruments, namely the Receptive Knowledge Test and the Productive Knowledge Test, to evaluate the students' knowledge of English adjective + noun collocations. It should be noted that the same list of 20 collocations, including 10 L1-L2 congruent collocations and 10 incongruent ones, was used to create the test items for both tests to ensure the comparability of the two tests. It was also ensured that these 20 collocations had yet to be taught in the grammar lessons.

The Productive Knowledge Test comprised 20 gap-filling items, each requiring the participants to complete a sentence by using a correct adjective to be combined with the given noun. This test had two sub-tests, Sub-test 1P, which included 10 test items about congruent collocations, and Sub-test 2P, which measured the students' knowledge of incongruent collocations. Since the

whole test aims to assess the students' collocational knowledge rather than their language competence, all contexts and missing word combinations were translated into Vietnamese to ensure that any errors students made were due to their lack of collocational knowledge rather than their failure to understand the given contexts. Following are the two test items extracted from the Productive Knowledge Test, the first of which is an example of congruent collocations (*quick meal*), and the second is an instance of incongruent collocations (*turning point*).

3. Thức ăn thừa có thể được hâm lại để anh có một **bữa ăn nhanh chóng** trước khi ra khỏi cửa.

→ The leftovers can be reheated for a _____ **meal** before you head out the door.

20. Có con là một **bước ngoặt** trong cuộc đời của họ.

→ Having the baby was a _____ **point** in their lives.

In addition to the Productive Knowledge Test, this study also incorporated the data obtained from the paragraph-writing tests for a more comprehensive evaluation of the learners' ability to use English collocations in different contexts. These writing tests were, in fact, part of the mid-term and end-of-term tests of the grammar course. Each student was required to write one paragraph of 150-200 words for each test's two topics. The combination of the Productive Knowledge Test and the paragraph-writing tests ensured the variety of the qualitative data, ranging from the adjectives supplied to complete the given word combinations to the adjective + noun collocations that students used naturally and spontaneously in their writing.

In the Receptive Knowledge Test, there was a total of 20 multiple-choice items, including 10 test items about congruent collocations (Sub-test 1R) and the other 10 about incongruent ones (Sub-test 2R). The inclusion of the same collocations in the Receptive Knowledge Test was anticipated to give a more comprehensive picture of the students' collocational knowledge. The reason is that, in the Productive Knowledge Test, the students might have written only the collocations they were sure of and avoided giving the responses they felt doubtful about, so the results from the Productive Knowledge Test only showed that the students knew which collocations were correct in the given contexts, but this does not mean that they knew which ones were incorrect. Meanwhile, receptive knowledge indicates the ability to recognize correct and incorrect word combinations (Gyllstad, 2007), so each test item in the Receptive Knowledge Test in this study consists of three options, from which the students had to select the best answer. Two sample questions from the Receptive Knowledge Test are provided below.

3. A. *quick meal* B. *fast meal* C. *Both are correct*

20. A. *turning point* B. *crossing step* C. *Both are correct*

For each question, there are three options, including the correct collocation, another word combination which is either incorrect or correct with the same meaning as the target collocation, and the last option, i.e., *Both are correct*. The incorrect word combinations were either derived directly from their Vietnamese equivalents, modified by replacing one or two elements of the right collocation with its/their synonym(s), or based on the students' inaccurate responses in the Productive Knowledge Test.

3.3 Data collection and analysis procedure

The data collection process involves the following steps. The Productive Knowledge Test was first administered to the students to prevent the students from remembering the answers from the Receptive Knowledge Test to fill in the gaps in the Productive Knowledge Test. After 30 minutes, the researcher gathered all the students' answer sheets before delivering the Receptive Knowledge Test. The students had 15 minutes to complete this test because it required significantly less time than the previous one. The total testing time was 45 minutes.

The Receptive Knowledge Test and the Productive Knowledge Test had a maximum score of 20. The responses to each question in this test were dichotomously coded in SPSS as either correct (1) or incorrect (0). In terms of the paragraph-writing tests, no predetermined score was specified. Instead, the students were encouraged to use several adjective + noun collocations in their writing. Each accurate collocation equals one point; thus, their score in the writing tests was the sum of the points they earned for providing the proper collocations.

With the aid of COCA, the two criteria of FREQ and MI (as specified in the operational definition of collocation for this study) were used to ascertain whether a combination of words used in the students' productive tests is a conventional collocation. Regarding the paragraph-writing tests, the L1-L2 congruency of the word combinations was determined after the collocations used in the students' paragraphs were collected. During this process, the back-translation technique was adopted with the support of reliable Vietnamese-English dictionaries to compare the English collocations with their translation equivalents in Vietnamese to determine whether a collocation is congruent or not and to avoid the researcher's subjective judgments. To determine whether Vietnamese learners recognized and used congruent adjective + noun collocations more accurately than incongruent ones, paired-samples *t*-Tests were conducted on the mean scores of the two subtests from the Receptive Knowledge Test and the Productive Knowledge Test. The frequencies of the congruent and incongruent collocations, including both correct and incorrect adjective + noun combinations, were also calculated and analysed for a complete understanding of L2 learners' production of collocations in a more natural writing task.

4. Results and Discussion

4.1 The role of congruency in the recognition of collocations

The first research question focuses on the role of congruency in Vietnamese students' recognition of English collocations. A paired-sample *t*-Test was used to compare the mean scores of the two sub-tests in the Receptive Knowledge test, the results of which are summarized as follows.

Table 1: Descriptive statistics of the results from the two Recognition sub-tests

	Mean	N	Std. Deviation	Std. Error Mean
Sub-test 1R (Congruent collocations)	4.903	72	1.4453	.1703
Sub-test 2R (Incongruent collocations)	5.708	72	1.9676	.2319

Table 2: Pair samples correlations regarding the scores of the Recognition sub-tests

	N	Correlation	Sig.
Sub-test 1R – Sub-test 2R	72	.208	.080

Table 3: Results of the paired-Samples *t*-Test for the scores of the Recognition sub-tests

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Sub-test 1R – Sub-test 2R	-.8056	2.1860	.2576	-1.3192	-.2919	-3.127	71	.003

Table 2 shows that the scores of the congruent collocation sub-test (1R) and the incongruent collocation sub-test (2R) were not correlated ($r = .208, p > 0.05$). This finding indicates that the learners' broad receptive knowledge of, for instance, incongruent collocations does not necessarily mean they are also good at recognizing correct congruent ones and vice versa. Moreover, Tables 1 and 3 show that there was a statistically significant difference ($t_{71} = -3.127, p < 0.05$) between the students' receptive knowledge of congruent collocations ($M = 4.903, SD = 1.4453$) and incongruent ones ($M = 5.708, SD = 1.9676$). The most striking finding from these data is that it is the conventional, congruent collocations that Vietnamese learners found more problematic to recognize.

The results from the Receptive Knowledge Test in this study are in contradiction with several previous studies (El-Dakhs et al., 2020; Jiang, 2022; Patekar & Košuta, 2022; Peters, 2015; Wolter & Gyllstad, 2011; Yamashita & Jiang, 2010; Yigit, 2021), which reported that L2 learners recognized the correct congruent collocations much better than the incongruent ones due to positive language transfer. Possible explanations for such a discrepancy in research results could lie in the collocation type selected for the test and how congruent and incongruent collocations may be taught and learnt. Specifically, the Receptive Knowledge Test in this study worked exclusively with adjective + noun collocations, whereas the tests in the previous studies mainly focused on verb + noun collocations. Different collocation types may lead to varying results in the role of congruency in L2 learners' collocation acquisition. In addition, how collocations are taught and learnt is also of great significance. According to Schmidt (1990), "noticing is the necessary and sufficient condition for converting input to intake" (p. 129). Typically, incongruent collocations receive far more attention from both teachers and students. The reason may be that these L2 collocations do not have any word-for-word translation equivalents in L1 and, therefore, are expected to cause more problems to L2 learners, which is known as negative language transfer.

As a consequence, special attention to incongruent collocations and efforts to learn them result in relatively better retention and better uses of these phrases. Meanwhile, congruent collocations are frequently ignored when they are seen or heard by L2 learners, for these combinations are easy to understand, and one would not anticipate any difficulty that may occur from such L1-L2 similarities. In fact, the problem will not arise until the learners actually use them in speaking and writing or are challenged to choose the word that can collocate with the word given. Because there are word-for-word translation equivalents of the L2 collocations in L1, students tend to learn these expressions analytically and may not pay adequate attention to word co-occurring information. When they need to use these collocations, they tend to recall individual words with the aid of translation due to significant L1-L2 similarities and then combine these single words

into a multi-word phrase without considering the lexical co-occurrence relations. These may be possible reasons why the students in this study were more likely to make mistakes with congruent collocations.

Take Items 3 and 14 in the Receptive Knowledge Test as examples. The collocation *quick meal* was used for Item 3. Since the English collocation *quick meal* has a congruent equivalent in Vietnamese, i.e., *bữa ăn nhanh*, most Vietnamese students, when encountering this collocation in reading and listening, do not need to make any effort to understand and remember it, but still, they could easily identify *quick meal* as a correct collocation when asked. However, since *quick* and *fast* are synonymous, the introduction of the lexical combination *fast meal* as the second option and *Both are correct* as the third option might have made the learners feel uncertain about the best answer for the test item. As a result, nearly 80% of the students failed to select the right answer to this question. However, this is not the case with Item 14 in the same test. The collocation *soft drink*, literally translated into Vietnamese as *đồ uống mềm*, has a Vietnamese equivalent, i.e., *nước ngọt* (**sweet water*). Because this set phrase is incongruent between L1 and L2, the students are more inclined to take notice of it and attempt to learn it by rote, and teachers also tend to draw students' attention to this phrase and remind them of it when they have a chance. Besides, breaking this incongruent collocation down into constituent words, e.g., *soft* and *drink*, may not help L2 students better remember this phrase, so students may choose to process and store it holistically, i.e., as a whole, and, therefore, they are less prone to compositionality and less likely to make mistakes. Perhaps this is why slightly over 75% of the participants gave the correct answer to Item 14 about *soft drink*.

In brief, the students were more susceptible to mistakes with congruent collocations due to their underestimation of the difficulties of the congruent collocations, their analytical approach to these multi-word expressions, and their lack of attention to co-occurrence information.

4.2 The role of congruency in the production of collocations

The second research question is whether students used congruent collocations more accurately than incongruent ones. Another paired-samples *t*-Test was run on the mean scores of the two sub-tests in the Productive Knowledge Test, the results of which are shown below.

Table 4: Descriptive statistics of the Production sub-tests

	Mean	N	Std. Deviation	Std. Error Mean
Sub-test 1P (Congruent collocations)	5.819	72	1.2141	.1431
Sub-test 2P (Incongruent collocations)	6.319	72	1.4125	.1665

Table 5: Pair samples correlations regarding the scores of the Production sub-tests

	N	Correlation	Sig.
Sub-test 1P – Sub-test 2P	72	.297	.011

As shown in Table 5, the scores of the congruent collocation sub-test (1P) and the incongruent collocation sub-test (2P) have a statistically significant, positive relationship ($r = .297, p < .05$). This suggests that a more accurate use of congruent collocations is associated with more accurate use of incongruent collocations, though the magnitude of this correlation is approximately weak ($.2 < |r| < .4$).

Table 6: Results of the Paired-Samples *t*-Test for the scores of the Production sub-tests

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Sub-test 1P – Sub-test 2P	-.5000	1.5655	.1845	-.8679	-.1321	-2.710	71	.008

More importantly, Tables 4 and 6 show that students' productive knowledge of congruent collocations ($M = 5.819$, $SD = 1.2141$) was significantly different ($t_{71} = -2.710$, $p < .05$) from their knowledge of incongruent collocations ($M = 6.319$, $SD = 1.4125$). Similar to the Receptive Knowledge Test results, Vietnamese learners tended to have considerably more difficulties in using congruent collocations than in using incongruent ones in a more controlled productive task (i.e., gap-fill test items).

Similar to the Receptive Knowledge Test results, congruent collocations are more likely to be at fault than incongruent ones in the Productive Knowledge Test. This finding confirms the result of El-Dakhs et al. (2020) and is consistent with Men (2018) and Cao and Badger (2021), although the last two studies used another research instrument (i.e., academic essays) to collect data. Therefore, the assertion that students are more susceptible to errors related to congruent collocations appears to be better grounded when justified and confirmed in different studies with different research instruments.

According to the researcher's observation and informal interviews with the participants, the students struggled with both congruent and incongruent collocations in the Productive Knowledge Test. Contrary to common belief, the students spent relatively much time on questions about congruent collocations, and two of the most problematic expressions can be found in Items 7 (*chân giả*, which is *artificial leg* or *fake leg* in English) and 10 (*tiền giả*, with *counterfeit money* as its English equivalent). In this test, the students had to translate the target phrases from Vietnamese to English, so they were more likely to be influenced by L1. Both Vietnamese expressions in Items 7 and 10 include the adjective *giả*, which can be translated literally to *fake*, *artificial*, *false*, *counterfeit*, etc. Because this word has many equivalents in English, the students felt hesitant when they were required to choose only one suitable adjective to fill in each blank. More importantly, after being provided with the right answers and further elaborations from the researcher, the students still showed a lack of confidence in their ability to use these conventional collocations correctly, complaining that it is hard to remember that they can say *artificial leg*, *fake leg*, *false nose*, *fake nose*, *false eyelashes*, *fake eyelashes*, *artificial eyelashes*, *counterfeit money*, and *fake money*, but not **false leg*, **artificial nose*, **false money*, or **artificial money*. Unlike these congruent collocations, the students did not face such a problem with incongruent ones after being given the answers and explanations. Although they might also have problems with incongruent collocations such as *small talk* (Item 13), *narrow escape* (Item 16), or *turning point* (Item 20) during the testing time, they found these combinations clearer and easier to remember after learning about them after the tests. This phenomenon may be due to the students' tendency to approach congruent collocations analytically and incongruent collocations holistically due to the nature of each collocation group.

The students' paragraphs on different topics written in the mid-term and end-of-term tests were also utilised as an adjective + noun collocation resource to examine how the students' actual use of English collocations in a freer writing activity. The results are presented as follows.

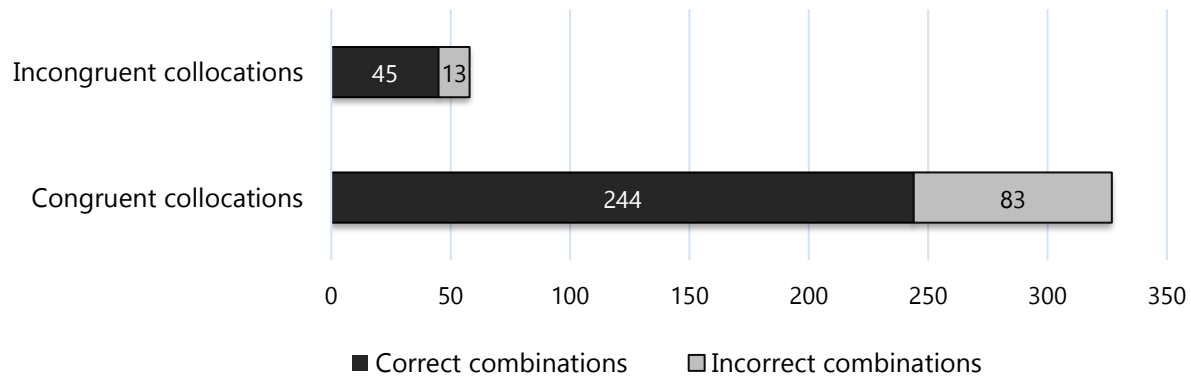


Figure 1: The frequencies of congruent and incongruent collocations, including both correct and incorrect ones, in the writing tests

As shown in Figure 1, the students tended to use far more congruent adjective + noun collocations (327 cases) than incongruent ones (58 cases) in their writings; specifically, the use of congruent collocations was more than five times as much as that of incongruent ones. As congruent collocations show great similarities between L1 and L2, the students might have felt more confident when using them in writing. Besides, several students at the beginning or intermediate levels tend to, and are even encouraged to, generate ideas in their L1 during the pre-writing stage (Friedlander, 1990; Lally, 2000); the students in this study might have used this strategy, which could be followed by their literal translation from L1 to L2. These might be why congruent collocations were more likely to appear in the students' writings. Unlike congruent collocations, the students seemed to be more careful when using incongruent collocations due to their L1-L2 differences and even avoided using them if they had any uncertainty about these combinations to minimize the mistakes and improve their performance. In second language acquisition, such a phenomenon is referred to as "avoidance strategy" or "avoidance behavior" (Kleinmann, 1977; Brown, 2000). This may explain why a very limited number of incongruent collocations were found in the students' paragraphs (58 cases).

Table 7: Descriptive statistics of the collocations used in the writing tests

	Congruent collocations		Incongruent collocations	
	FREQ	%	FREQ	%
Deviant collocations	83	25.38	13	22.41
Well-formed collocations	244	74.62	45	77.59
Total	327	100.00	58	100.00

As can be seen in Table 7, the percentage of deviant incongruent collocations was 22.41%, which was almost 3% lower than the proportion of deviant congruent collocations. There may be several reasons for this. Firstly, because of the students' avoidance strategy and cautious attitude, incongruent collocations occupied merely 15% of all the attempts to use collocations in the paragraphs, and the rate of errors related to incongruent collocations was also lower and less

varied. Secondly, the participants in this study were at intermediate and upper-intermediate levels of English; therefore, they had sufficient exposure to English and could have encountered or learned the target incongruent collocations before. Due to the salient features of this collocation group and the learners' extra attention to those different from their L1 equivalents, the students were more likely to remember these collocations much better than the congruent ones. The last reason lies in the nature of some incongruent collocations, at least many of the phrases used in the students' paragraphs (e.g., *turning point*, *small talk*, *soft drink*, *intelligent life*, *early age*, *fresh water*, *dark chocolate*). Many of them are, to some extent, idiomatic, which means the constituent words in these collocations can hardly be substituted with their synonyms. Students at higher levels of English proficiency are usually aware of this phenomenon, so they will not risk changing the composition of these incongruent collocations. Moreover, the constituent words in many of these collocations (e.g., *turning point*, *soft drink*, *early age*, *dark chocolate*) do not have many synonyms, which lowers the risk of students replacing any of them with their synonyms, resulting in fewer errors with this collocation group.

The results from this study also indicate that using many congruent collocations does not necessarily mean that the students are proficient at using them. The evidence is that the deviant congruent collocations accounted for one-quarter of all the attempts to use congruent collocations in writing. This percentage may not be exceptionally large, but it is somewhat startling and disconcerting. The students made mistakes in what they felt confident about, and, as a result, they might not think about checking whether these expressions were deviant or not. For example, the majority of the student participants (slightly over 92%) selected both *bad quality* and *low quality* as conventional collocations (Item 8 in the Receptive Knowledge Test); however, *bad quality* is a deviant collocation, and approximately one-tenth of the participants did continue to use this incorrect word combination in their paragraphs written for the mid-term or end-of-term exams. This suggests that L2 students might not be aware of the mistakes until these very mistakes are highlighted and explained directly to them. Besides, as the participants in this study were at high levels of English, they tended to use unfamiliar words in their writing tests, which were also their mid-term and end-of-term tests for the Grammar course, to demonstrate their lexical range and variety to gain higher scores. However, many of them failed to consider the word co-occurrence of these uncommon phrases. For example, instead of using the familiar collocation *wonderful vacation*, Student 17 used *stunning vacation*, and Student 66 used *glamorous vacation* to express the same meaning. These combinations are unconventional and therefore considered deviant, though *wonderful*, *stunning*, and *glamorous* are synonyms in this context. Thus, using uncommon phrases by substituting one familiar word in the phrase with its synonym, which belongs to a more advanced-level wordlist may be counter-productive if students do not pay adequate attention to lexical co-occurrence.

5. Conclusion

Contrary to the theory of language transfer (Weinreich, 1953; Winford, 2003), the results from the Receptive Knowledge Test, Productive Knowledge Test, and paragraph-writing test indicate that the problems which incongruent collocations might cause to L2 learners are not as serious or as common as congruent ones. In other words, L2 students are even more likely to recognize and

use incongruent collocations more accurately than congruent ones. Therefore, similarities between L1 and L2 collocations do not necessarily lead to positive language transfer, especially in the case of collocations, and differences between L1 and L2 are not always an obstacle to L2 learners' collocation acquisition.

Due to the possible problems that L2 learners may have with both congruent and incongruent collocations, as shown in this study, EFL teachers and students should not only pay attention to incongruent collocations due to their salient differences; in fact, more effort should be made to learn congruent collocations because the problems they cause are less noticeable but more common and more difficult to address than incongruent ones even after well-formed congruent collocations have been taught to the students. Hence, it is advisable that teachers raise students' awareness of the difficulties that congruent collocations may cause and provide them with more practice in collocations, especially the confusing ones. As the results from the correlation tests in this study show, there was a weak relationship or even no relationship between the learners' knowledge of congruent and incongruent correlations, there should be no assumption that students only need to focus on one collocation group, e.g., incongruent collocations, and then their knowledge of the other group, e.g., congruent collocations, would automatically increase. It should be highlighted that both congruent and incongruent collocations should receive adequate attention from both teachers and students. The use of corpus and collocation dictionaries should also be taught to the students to help them communicate their ideas more effectively and efficiently, especially to improve learner autonomy inside and outside the classroom.

Due to time and resource constraints, this study has certain limitations with respect to the sample size and sampling method. To enhance its generalizability, future research could collect a larger sample of students selected through probability sampling. Moreover, the results of this study indicate several avenues for further research, including (1) the analysis of collocational errors made by EFL learners of different mother tongues, (2) the examination of the development of receptive and productive knowledge of collocations over time, and (3) the investigation into the acquisition of congruent and non-congruent collocations among EFL learners at different levels of English proficiency.

Declaration of Conflicts of Interests

The author(s) declared no potential conflicts of interest.

References

- Benson, M., Benson, E., & Ison, R. (2010). *The BBI combinatory dictionary of English: Your guide to collocations and grammar* (3rd ed.). Benjamins.
- Brown, H. (2000). *Principles of language learning and teaching* (pp. 122–130). Longman.
- Cao, D. & Deignan, A. (2019). Using an online collocation dictionary to support learners' L2 writing. In C. Wright, L. Harvey & J. Simpson (Eds.), *Voices and practices in applied linguistics: Diversifying a discipline* (pp. 233–249). White Rose University Press.
- Cao, D. & Badger, R. (2021). Cross-linguistic influence on the use of L2 collocations: the case of Vietnamese learners. *Applied Linguistics Review* 2021, 1-26. <https://doi.org/10.1515/applirev-2020-0035>

- Celce-Murcia, M. (2008). Rethinking the role of communicative competence in language teaching. In E. A. Soler & M. P. S. Jordà (Eds.), *Intercultural language use and language learning* (pp. 41–57). Springer.
- Chorbwhan, R., & McLellan, J. (2016). First language transfer and the acquisition of English collocations by Thai learners. *Southeast Asia: A Multidisciplinary Journal*, 16, 16–27.
<https://fass.ubd.edu.bn/SEA/vol16/SEA-v16-chorbwhan.pdf>
- Clear, J. (1993). From Firth principles: Computational tools for the study of collocation. In M. Baker, G. Francis & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair*. John Benjamins.
- Cowie, A. P. (1981). The treatment of collocations and idioms in learners' dictionaries. *Applied Linguistics*, 2(3), 223–235. <https://doi.org/10.1093/applin/II.3.223>
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, 20(1), 29–62.
<https://doi.org/10.1515/text.1.2000.20.1.29>
- Firth, J. R. (1957). *Papers in Linguistics 1934-1951* (pp. 190–215). Oxford University Press.
- Foster, P. (2001). Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 75–94). Longman.
- Friedlander, A. (1990). Composing in English: Effects of a first language on writing in English as a second language. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 109–25). Cambridge University Press.
- Granger, S. (2018). Formulaic sequences in learner corpora: Collocations and lexical bundles. In A. Siyanova-Chanturia & A. Pellicer-Sanchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 228–247). Routledge.
- Gui, S. C. & Yang, H.Z. (2003). *Chinese learner English corpus*. Shanghai Foreign Language Education Press.
- Gyllstad, H. (2007). *Testing English collocations: Developing receptive tests for use with advanced Swedish learners* [Doctoral dissertation, Lund University]. Lund University Publications.
<https://lup.lub.lu.se/search/ws/files/5893676/2172422.pdf>
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Routledge.
- Hong, A. L., Rahim, H. A., Hua, T. K., & Salehuddin, K. (2011). Collocations in Malaysian English learners' writing: A corpus-based error analysis. *3L: The Southeast Asian Journal of English Studies*, 17 (Special Issue), 31–44.
- Howarth, P. (1998). The phraseology of learners' academic writing. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 161–186). Clarendon Press.
- Kleinmann, H. H. (1977). Avoidance behavior in adult second language acquisition. *Language Learning*, 27(1), 93–107. <https://doi.org/10.1111/j.1467-1770.1977.tb00294.x>
- Kurosaki, S. (2012). *An analysis of the knowledge and use of English collocations by French and Japanese learners* [Doctoral dissertation, University of London].
<https://core.ac.uk/download/pdf/28902137.pdf>
- Lally, C. G. (2000). First language influences in second language composition: The effect of pre-writing. *Foreign Language Annals*, 33(4), 428–432. <https://doi.org/10.1111/j.1944-9720.2000.tb00623.x>
- Laufer, B. (2010). The contribution of dictionary use to the production and retention of collocations in a second language. *International Journal of Lexicography*, 24(1), 29–49.
<https://doi.org/10.1093/ijl/ecq039>
- Laufer, B. & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61(2), 647–672. <https://doi.org/10.1111/j.1467-9922.2010.00621.x>
- Leech, G. (1974). *Semantics: The study of meaning*. Penguin.
- Lewis, M. (2000). *Implementing the Lexical Approach: Putting theory into practice*. Language Teaching Publications.

- Men, H. (2018). *Vocabulary increase and collocation learning: A corpus-based cross-sectional study of Chinese learners of English*. Springer.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223–242. <https://doi.org/10.1093/applin/24.2.223>
- Nesselhauf, N. (2005). *Collocations in a learner corpus*. Benjamins.
- Patekar, J., & Košuta, N. (2022). Croatian EFL learners' collocational competence: Congruent and incongruent collocations. *Training, Language and Culture*, 6(3), 9–31. <https://doi.org/10.22363/2521-442X-2022-6-3-9-31>
- Peters, E. (2015). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research*, 20(1), 113–138. <https://doi.org/10.1177/1362168814568131>
- Schmitt, N. (2012). Formulaic language and collocation. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1–10). Blackwell.
- Shin, D. & Chon, Y. V. (2019). A multi-word unit analysis: COCA multi-word unit list 20 and ColloGram. *The Journal of Asia TEFL*, 16(2), 608–623. <https://doi.org/10.18823/asiatefl.2019.16.2.11.608>
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- Wolter, B. (2006). Lexical network structures and L2 vocabulary acquisition: The role of L1 lexical/conceptual knowledge. *Applied Linguistics*, 27(4), 741–747. <https://doi.org/10.1093/applin/aml036>
- Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics*, 32(4), 430–449. <https://doi.org/10.1093/applin/amr011>
- Wolter, B. & Yamashita, J. (2015). Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics*, 36, 1193–1221. <https://doi.org/10.1017/S0142716414000113>
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.
- Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly*, 44(4), 647–668. <https://doi.org/10.5054/tq.2010.235998>
- Yigit, G. (2021). The influence of L1 on the acquisition of L2 collocations: Turkish ELT students. *TESL-EJ*, 25(3), 1–13. <https://tesl-ej.org/wordpress/issues/volume25/ej99x/ej99a2/>

Ngoc Thai Bao Pham received a Master's degree in TESOL from the University of Social Sciences and Humanities, Vietnam National University – Ho Chi Minh City (USSH, VNU-HCMC) in 2016 and is currently a Ph.D. candidate in contrastive linguistics at the same university. She has taught English to Vietnamese students for over a decade and is presently a lecturer at the Faculty of English Linguistics and Literature at the USSH, VNU-HCMC. Her research interests include English language teaching, cognitive linguistics, and contrastive linguistics.
